<<<Back to Table of Contents

On to Epilog (or prolog)>>>

Geoscience after IT: Part M

Business requirements drive the information system, and provide coherent frameworks

T. V. Loudon
British Geological Survey, West Mains Road, Edinburgh EH9 3LA, U.K.
e-mail: v.loudon@bgs.ac.uk

Postprint of article in Computers & Geosciences, 26 (3A) April 2000, pp. A123-A132

Abstract - The roles of participants in the information system are changing, and this is reflected in their business groupings and motivation. IT brings greater flexibility to the record, but without a coherent framework, cyberspace becomes a chaotic sludge of trivial ephemera. Cataloging and indexing, peer review by editorial boards and the disciplined approach of information communities can impose the necessary order and standards. Metadata and data models can help to maintain a clear structure for geoscience. Business aspects link the objectives of the investigator to the framework of the science, defining the logic of reorganization and providing incentives to drive the system.

Key Words - Catalogs, editorial boards, information communities, information system strategy, business aspects.

1. Activities, participants, roles and driving forces

Subsystems are selected to minimize their interactions (part I, section 2.2). Nevertheless, much of the interest lies at the interfaces. Scientific investigations are conducted at the interface between the real world and the information system, drawing information from the repositories, testing or extending it by observations and measurements in the real world, and returning with conclusions that may be added to the knowledge base. The scientists' activities (D 7) are usually described by verbs, such as investigate, integrate, explain, curate, communicate. During a project, there is at least one cycle of **activities** (applying processes to objects), such as plan, undertake desk studies and field work, analyze, report, review, possibly return to additional study of the literature, more field work, and so on. Fig. 1 shows them within a circle, to avoid an arbitrary beginning or end.

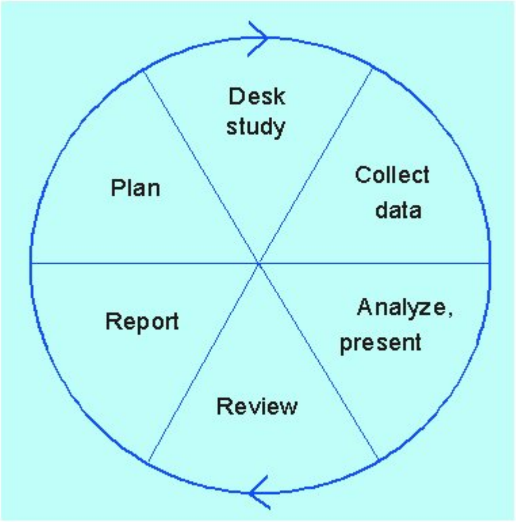

Fig. 1. Some activities in a project. A cycle of activities that may be followed more than once during an investigation.

The investigators and the users of the information interact with many other **participants** (those playing a part in the operation of the information system). Managers may define the business setting, possibly standing in as proxies for other stakeholders such as customers or stockholders. In an educational environment, supervisors may interpret the views of professors and other academics. Contact with the recorded knowledge base is likely to be through **intermediaries** (who assist the user or contributor to interact with some aspect of the system), such as librarians, information scientists, booksellers and curators. Collection of data may be assisted by laboratory staff, instrumentation experts, field and laboratory assistants. Recording the results may involve typists, data-entry specialists, reviewers, editors, referees, curators, catalogers, database administrators, printers and publishers. The managers or supervisors are likely at all stages to advise, monitor progress, and ensure that the results conform to the intended objectives. The recommendations of standards organizations have a bearing at all stages. In addition, scientists interact informally with others working in a similar area or topic, with customers, and with those who can supply more detailed information or who are involved in broader studies.

Technology reduces the dependence of authors and users on intermediaries, such as those just mentioned (see also B 1, M 2.1). As the information industry changes in response to new technology, the new participants tend to be described in terms of their roles, and some of the old categories can be merged with the new. The roles include clients; users; information owners, keepers and suppliers; database and repository managers; Internet service providers; network (communications and delivery) operators; webmasters; application developers; standards setters and quality assessors. The client/server mode of operation places responsibility for storing and serving the information with the originator, probably delegated to a proxy, and the form of presentation with the client who reads the information. The considerable support from the IT industry and the service suppliers, the / between client and server, is a vital, often neglected, component, but is not our subject here.

The participants work together in **business groupings**, such as oil companies or academic faculties, sharing broad objectives and ways of working. The organization is likely to be staffed with a range of experts to meet the demands of the projects it undertakes. It probably provides financial support and shared facilities, such as access to records, libraries, laboratories and computing facilities. The participants are themselves likely to be represented as objects in other databases, such as staff lists and personnel files. These may well contain information useful to the scientist, such as an e-mail address or the name of a student's supervisor. The need to communicate crosses all boundaries and cannot be limited by system definitions.

Information is driven through the system by powerful forces. Curiosity or commercial gain may be the initial incentive for investigation. A desire to share the results, perhaps to gain kudos, promotion or scientific standing, can carry the process forward. Without such forces, it is unlikely that the system would operate at all, and their identification (preferably without undue cynicism) is an important part of analyzing the system. The **motivating factors** that drive the participants (Herzberg et al, 1993) include opportunities for achievement, recognition and advancement, and the chance to exercise creativity and take on responsibility. In any change to the system, motivation must be kept in mind to ensure cooperation in the new development. Its form may determine, for example, whether information sharing or information hoarding seems more attractive (I 8.2, M 3.2). The motives of, say, the scientist and the publisher may conflict, creating tension within the information system that has to be managed by negotiation. The information system is a social creation involving many disparate contributors in a shared activity. It will succeed or fail (Peuquet and Bacastow, 1991) depending on the motivation of all concerned.

2. Frameworks for models

Recorded information has in the past been formalized and fragmented into maps, reports, databases, archives and collections. We are now at an early stage in the global development of the hypermedia knowledge repository also known as cyberspace. Contributors of information are not constrained by the form of final presentation. This can be decided by users to meet their specific requirements. Users can bring together information from many sources, align differing ideas, employ visualization techniques, and present the results as they see fit. However, although the system can accommodate individual contributions on their own terms, that merely transfers to users the task of finding and integrating information from different sources. Flexibility is obtained at the cost of a clear structure.

From the desktop we reach a vast pool of diverse knowledge. Unfortunately, it seems at times like a chaotic swirling sludge of trivial ephemera. Metadata and the associated standards are essential to sharing knowledge, but do not in themselves provide the framework for organized thought. We can understand words without recognizing a coherent story. To bring order to this chaos we need a framework that reflects the structures of our conceptual models.

Projects generally aim to gather new information or develop new ideas. Again, a clear structure might make the process more efficient. In reporting the results, there should be no need to repeat large amounts of what has already been recorded. Instead, linkages should as far as possible give an indication of the dependency on earlier work and earlier ideas. This can be done by the author in the course of preparing the new document. Given a well-structured context, it should be possible to indicate precisely what has been assumed and where the ideas depart from those generally accepted. This would help to assess knock-on effects when ideas or definitions change. It could also indicate to the reader, in a detailed and objective way, the level of knowledge needed to assimilate the new ideas contained in the document, and offer precise guidance on where additional background can be found. It should not be necessary to complete a project before making results available to others, as fragments can be linked in to the existing context, and new versions replace old ones as the work proceeds. Nevertheless, to some extent projects supersede previous studies and therefore cannot be rigidly constrained within a pre-existing framework.

Three parallel approaches to a more coherent framework come to mind.

1. One is cataloging. For example, bibliographic information for project documents could be held in a library catalog; or new hydrocarbons data could be recorded, placed in a repository and cataloged following POSC standards. The catalog provides structure and a means of access. The reputation of the data supplier or archive manager gives some assurance of quality.
2. A second approach, well suited to exploratory projects, is extension of the current scientific literature (I 6), to embrace new mechanisms of delivery while retaining the structure and evaluation imposed by the editorial board, as described in L 3. This implies that each document is largely self-contained and self-explanatory. As previously mentioned, archiving problems may arise with multimedia documents (L 6.3).
3. A third approach is for an information community to define a general model (M 2.3) for structuring relevant knowledge within the scope of their activities, treating individual investigations as subprojects within a unified framework that evolves as ideas develop.

The three approaches, considered in M 2.1 and 2.2, are not mutually exclusive. The same object could be relevant in more than one framework, and can readily be shared by hypertext reference. For example, the description of a fossil might be archived only once, but referenced from a geological survey model, an oil exploration repository and a paleontological journal. In this fast evolving field, it is likely that these and many other frameworks for shared knowledge will be explored. As members of the geoscience community, it is our task to drive this evolution - to understand and appraise, encourage or condemn.

2.1 Realigning responsibilities

Archiving has in the past been the responsibility of libraries. The copyright in the documents, however, is generally retained by their publishers, who could create an electronic archive of them as a source of significant future income. In due course, such archives would almost certainly improve the service and reduce the cost. Much of that saving would come from the reduced archiving role of the libraries. Libraries, however, are at present the main channel of government funds in purchasing conventional publications. The publishers may be reluctant, for the time being, to upset their main customers. In a travesty of the market place, publishers may even offer electronic copies of science journals at a higher price than the paper version on the grounds that even if they cost less to produce, they offer more to the purchaser.

If some of the functions of libraries are at risk from electronic archives, the same could be said for publishers (see Butler 1999). Authors may be tempted to publish their own papers. In the past, this would have condemned the paper to obscurity. In the future, cataloging information is likely to be part of the document or digital object, and thus the responsibility of the object's owner. It is retrievable by search engines, some of which, for efficiency, would copy the cataloging information into an index and possibly extend it. Readers could thus find and retrieve the paper independently. But there are problems. The quality of an unrefereed document has not been assessed, and there is no indication, other than the author's reputation and affiliation, of its scientific standing. There is therefore little kudos for the author, and no guidance for the reader about whether its findings are valid or significant. Furthermore, without some assurance that the paper will be widely available within a long-term archive, it cannot be seen as part of the permanent scientific record.

Solutions to these problems are in the hands of editors rather than publishers. Editors and referees are more likely to be concerned, as at present, with whether contributions deal with an appropriate topic at a suitable level, meet the house style and agreed standards, evaluating their relevance and quality, and ensuring that they are original, inoffensive, and give credit where it is due. Readers then have the task of assessing the 'brand names' of the editorial boards, just as they would expect today to have a view on the quality of a particular journal.

The role of the publisher, on the other hand, may be subsumed into repository or archive management, concerned principally with organizing and maintaining an information system and making its contents available. Some scientific societies have taken on the role of publisher, for example, the Institute of Physics (1999). In some fields, large commercial publishers may dominate, because of their financial resources, global reach, wide coverage of many disciplines, marketing skills, and above all their copyright of existing content (ScienceDirect, 1999). In the long run, costs must be recovered for access to an electronic repository, and the profit potential, particularly if charges are visible to the reader, may prove controversial.

Features which users might look for in such an information system (K 3) include:

- *stability* - there should be a clear and credible commitment to long-term preservation, access and maintenance of all information
- *usability* - provenance, ownership of intellectual property rights, and terms and conditions of use should be clear
- *fairness* - charges and conditions of use should be seen as fair, reasonable, competitive and consistent
- *reliability* - the user should be able to assess the accuracy and quality of all information
- *comprehensiveness* - within the demarcated scope of the service, the user should be confident that all likely sources are included or referenced, including the most recent work
- *convenience* - the system should be easy to use through a consistent, familiar interface, and provide a rapid and efficient response
- *clear structure* - available and relevant material should be easy to find and have pointers to related information

The Digital Object Identifier (L 3) is a basis for such a system, building on the existing scientific literature. The digital object is the scientific paper, supported by entries in indexes and catalogs. A consortium of publishers has taken an intitiative (ScienceDirect, 1999) to supplement, and potentially replace, paper copies with items archived electronically, for example in SGML. Versions for browsing and printing, for example in HTML, XML and PDF, are generated from the archive and accessible through the publisher's gateway, which controls access and imposes charges if required. A wide range of refereed literature, with indexes, abstracts and catalogs, can thus be made available from the desktop. The flexibility and ease of use of the Web browser complement the authority, structure and permanence of the scientific literature. The digital objects can of course be hyperlinked, and references reached by a click. As they share the same desktop interface, the formal literature can link to ephemera, detail, and work in progress recorded on the World Wide Web. Equally the literature can have links to and from the spatial models described in the next section.

The scientific paper of around 5000 words is a convenient length for downloading to the desktop, and appropriate for marshaling and presenting a coherent view of a specific argument. More extended accounts, such as topic reviews and books, are normally arranged in chapters dealing with separate aspects of the subject. The chapter, rather than the book, might be seen as suitable for cataloging as a retrievable unit, analogous to the scientific paper. Electronic archives should focus on the content not the container. Older distinctions based on the format of presentation, such as book, serial or reprint, are likely to blur. The general scientific literature, however, is likely to be archived as sets of discrete objects or documents, and this may be inappropriate for some of the tightly structured work of information communities.

2.2 The information communities

The OGIS Guide (L 4) points out that each scientist has a unique view of the world, and that this makes communication more difficult (Buehler and McKee, 1998). They identify **information communities** – collections of people who, at least part of the time, share a common world view, language, definitions and representations – and explore possible means of communicating between such groups. An example might be NOAA, the Department of Mines and the USGS, each with their own objectives, methods, terminology and standards. Understanding the concepts of an information community can be helped by a strong framework of data models with clearly defined terms and relationships.

Valid interpretation across community boundaries is likely to depend in large part on the background knowledge of the human interpreter. This is not available, nor likely to become available, to the computer. As Kent (1978) pointed out, language is a powerful tool to reconcile different viewpoints, and a basis for communicating background knowledge both of large concepts and of the details of a single object. Written explanations are therefore needed in close association with spatial and data objects at every level of detail.

We can already see in the World Wide Web the emergence of a global knowledge network, using hypermedia to express and relate ideas. There is a clear distinction between cross-references among objects, which call attention to some relationship or analogy, and the tightly linked conclusions that emerge from a project based on a single coherent set of background assumptions, objectives and working methods. Within the loose global linkages of the hypermedia knowledge repository there are more tightly organized structures of geoscience information managed by defined information communities.

Large information communities, such as geological surveys, already publish many of their own findings, and should find this easier in an IT environment. Their internal refereeing and quality assessment procedures, their copyright ownership, their brand name and reputation among their customers are already established. They should therefore be able to meet most of the users' criteria in 2.1.

As well as their own findings, a survey may hold data originally collected by external organizations for various purposes, such as site investigation or mineral assessment. The accuracy of the data is variable and cannot appropriately be judged by the survey. Provided this is made clear, however, and the source and ownership of data sets are clearly identified in the metadata, the user can evaluate them against the quality-assessed survey view of the same area, and vice versa. The survey is adding value by making the information available in context. There are benefits to the contributor in placing records within the structure of a repository where the costs of initial design, installation, marketing and maintenance are spread across many users. There are benefits to the repository in achieving more comprehensive coverage by accepting external contributions.

The requirement for up-to-date information seems to conflict with the need for a permanent record of earlier views. This can be overcome by archiving date-stamped previous versions, or by retaining the ability to reconstruct them from journalized changes, as generally only a small part of a document or data set is superseded. The task of maintaining versions should not be underestimated, for while it can be readily handled in a prototype, it could be the dominant issue in a production system (Newell et al, 1992).

In fast developing technology, the lead organization tends to keep moving ever further ahead of the pack, because users prefer a single mainstream solution. The leader can set standards while others inevitably fall behind. A **winner-take-all** situation develops, to be broken only by user dissatisfaction, by competitors using technology more effectively or catching a new wave of technological advance, by financial muscle, by political interference or a combination of these. Even within a small niche, such as a country's geology, users may prefer a single source of survey information. All users can then work on the same basis, and different areas can readily be compared.

On those grounds, a survey or similar organization (indeed any group dominating its niche and working to shared, comprehensive standards) can be well placed to maintain its market position. It just needs to stay in the forefront of technology, keep in line with changing standards, and satisfy customers and politicians. Because information technology bridges national and disciplinary boundaries, standards must be international and standards within geoscience must be consistent with those in related fields. Close collaboration with a range of other organizations is therefore essential. Some organizations can share information system resources through “extranet crossware” (Netscape, 1999). Both sides gain from the links (**win-win**), as well as customers benefiting from good service at reasonable cost.

An information community exists because its members share objectives and are organized to find an integrated solution. A geological survey (I 8.1) is an example of an information community, one of its roles being to assemble basic geological information about an area in a form which can be used in many other applications (rather than being collected separately for each project). The conventional means of achieving a coherent overview is to publish standard series of maps with accompanying memoirs and explanatory reports, all to consistent standards. As this is firmly embedded in old technology, surveys have had to review their work from first principles. The British Geological Survey, an example of a medium-sized survey, considered the issue of their basic geoscientific model (Ovadia and Loudon, 1993) as described in the next section.

2.3 A basic regional geoscience framework

The earlier description of the geoscience information system (I 2.3) gave some impression of its scope and form, but said little about its scientific content. The triangular image of increasing abstraction in L, Fig. 3 hints that there is some shared, general model – the paradigm that geoscientists have at the back of their minds. If so, there should be a route from a single set of observations at the bottom of the triangle, such as a soil profile, linking upwards at higher levels of generality through the entire body of existing knowledge. Indeed, the claim to be a science suggests that the body of knowledge should be coherent and internally consistent. A greatly simplified overview is required to provide an overall structure into which observations and ideas can be fitted, and from which relevant information can be retrieved.

The framework of a general geoscience model can help to bring order to a multitude of investigations whose varied business aims lead to a diversity of approaches. A single, coherent, general model can specifically address the area of overlap and thus help to avoid unnecessary duplication. The task of developing that model and sharing the results can appropriately be assigned to an identified information community, such as a geological survey. The need for cooperation with related information communities is illustrated by, for example, the links between topographic and geological mapping.

Geological, topographic and related surveys worldwide have developed such models of national aspects of geoscience. Examples can be found in Australia (Australian Geodynamics Cooperative Research Centre, 2000), France (BRGM, 2000), Canada (Lithoprobe, 2000) and the United Kingdom (Adlam et al, 1988). Their concern is to convey knowledge of the consequence of geological and related processes, states and events in geological time and space. Their findings, which have a strong spatial element, have generally been expressed as maps and reports on specific areas. Geological maps may list the various rock units present in the area (**classification** and **nomenclature**), and by relating their location to a topographic base map, show their spatial distribution (**disposition**) at or near the earth's surface. Drift and soil maps may show the disposition of sequences of units. Orientation measurements, intersections with the topography, and cross-sections give an impression of the three-dimensional form, sequence, shape and structure (**configuration**) of the units.

Generalized sections and text comments give an indication of the lithological and petrographical **composition** of the material. Specialist maps might give information about the geochemical composition of the material, the geophysical **properties** of the rock mass or the geotechnical properties of individual units. Paleogeographical maps and palinspastic reconstructions can be used to express a view on their formation and **historical development**. Symbols on the map may show wells, traverses, measurement stations, outcrops, and collection localities as points where **evidence** was gathered to support the conclusions.

Many aspects, such as detailed descriptions and accounts of processes, can be addressed more satisfactorily in text than on a map. The paleontologist studying a single fossil, or the seismologist studying an earthquake, may indeed consider a general geoscience model to be irrelevant. But their findings are ultimately related to some framework of space and time, viewing the fossil as a component of the material of a rock unit, throwing light on its history; and the earthquake as an event resulting from the reaction of the material to its properties and stress history.

Reports and maps are often closely associated, but perhaps maps give clearer pointers to a general model, because their graphical symbolism, uniformity, and the need for worldwide coverage require a formalized and consistent approach. However, the conventional map is a product of a particular technology. We are looking for an underlying model which refers to the scientific concepts, not the technical solutions, for our interest is in how technology can evolve to fit scientific needs (see Laxton and Becken, 1995). The concepts must be as free as possible from their form of presentation.

In a **general geoscience spatial model**, the objects of interest are the earth and parts of the earth, such as rock units or their bounding surfaces. The aspects of interest just mentioned are their disposition, configuration, composition, properties, history and evidence.

The underlying concepts are familiar. They address issues analogous to those that might worry a three-year-old child on looking into a dark room.

- What is in there and what is it called? (Object classification and nomenclature)
- Where is it? (Disposition)
- What does it look like? (Configuration)
- What is it made of? (Composition)
- What does it do? (Properties)
- How did it get there? (History and geological processes)
- How do I know? (Evidence and business aspects)

We try to develop and convey the knowledge (held in our brains) of states, processes and events, sequenced in time and patterned in space, which we believe may account for our observations within our accepted world view. The types of model with which geoscientists are concerned largely determine the unique characteristics of their information system. In particular, the spatial model (G 2) is the key, not only to the disposition and configuration of objects, but also to understanding many of the relationships of their composition, properties and processes. IT may offer radical improvements in implementing the framework.

Where a strong framework and good retrieval techniques are in place, the survey model can tie into contributions from external projects, like a commentator adding footnotes to an existing story. Ideally, it should support interwoven stories dealing with any relevant topic, tied to geological space and time and the object-class hierarchies defined in the metadata. Data models define the scope and relationships of the topics considered, and provide a structure for storage and retrieval of information. The content may be complete for all the subject areas and topics, although level of detail and date of last revision will inevitably vary. Here is a context where contributions can be evaluated, stored and found when required, and a means of reconciling information obtained for differing business purposes. Conflicting and changing views can be held side by side, for evaluation by users.

Models such as this can provide firmly structured areas embedded in the more flexible hypermedia knowledge repository. Some information communities, such as oil companies, have more clearly defined business requirements, and precise ideas about the geoscience information required to support them. Academic studies, in contrast, may have fewer preconditions, and a need to follow ideas wherever they may lead. They must choose different points on a trade-off. On the one hand, well-defined structures and consistent standards bring reduced redundancy, increased relevance and efficient access. On the other hand, the scientific literature offers greater flexibility and ability to cope with change. The cost is greater repetition and greater effort to comprehend the diversity of ideas and modes of expression. The scientific literature is already evolving to offer hypermedia documents within distinct topic areas overseen by editorial boards.

We thus see the development of a flexible hypermedia knowledge repository. Within it, structured areas are provided by information communities of all sizes and forms, from individual businesses to consortiums of business partners, geological surveys, editorial boards for geoscience literature, and the organizations that help to establish, formalize and encourage the use of standards.

3. Business aspects

Any geoscience project is embedded in some kind of business - mineral development, civil engineering, survey, education, research, or whatever - which sets its objectives, resources and time scale. All the information systems that deliver information for the business, including the geoscience information system, are changing because of IT. Most businesses follow a yearly cycle of reviewing progress, deciding priorities, allocating funds and so on, according to a business plan. Feeding information into this, and therefore synchronized with it, there may be an **information system strategy** which supports the business objectives (CCTA, 1989). It sets out a plan (for the various parts of the business) for development of the information systems, policies, programs of work and IT infrastructure. While the strategy may be the responsibility of an IT department, geoscience needs must be taken into account and fed through to the business plan at the appropriate time. The geoscience manager who wishes to take full advantage of IT must therefore keep the business aspects in mind.

The unpredictable course of technology will itself respond to business needs. Views on mainstream developments in IT can be culled from the Web pages of the major software suppliers, such as Microsoft, Oracle and Sun (their Web addresses can be found by inserting their names between www. and .com). Those with a more academic approach may be more interested in open source software, such as Linux or GNU (place between www. and .org for their Web addresses). Such code can be amended for specific applications and much of the software is free. Today's standard procedures may be by-passed by tomorrow's technology, and planning should therefore be flexible and kept under continual review.

3.1 Organizational consequences

In areas such as word processing, computer-aided design and Web searching, computers can assist users to carry out tasks which otherwise would require, say, a typing pool, publisher, drawing office, and library. Some changes to roles in the organization were considered in M 1. In general, intermediaries between the originator and the user of information can either be eliminated (**disintermediation**), or given a changed role, for example in providing advice on design and layout or development of standards, or in providing information systems maintenance and training.

Computer support can assist project planning and monitoring. Because computer-mediated information can be made available rapidly and widely within an organization, employees can respond to plans and requirements within a less complex management hierarchy (**delayering**) and with greater independence of individuals and groups.

Rather than regarding collection and management of information as closely linked activities, with data collectors responsible for looking after their own results, **standards** provide flexibility to combine or separate the responsibilities as appropriate. Information can be maintained by the originator during the course of a project and still be available to others over network links. Without necessarily altering the standard format of the information, it can in due course be passed to the control of a repository for long-term security.

Large amounts of data can be analyzed by computer provided they meet uniform standards. Where detailed standards are in place, many groups from many organizations can contribute shareable information. Data can be stored and managed in a shared repository. For example, POSC (L 5) has assembled standards that enable data from many sources to be shared through local and international repositories, where the task of data management is handed over (**outsourced**) to specialists. The result is huge savings to individual companies, and generally more reliable access to information.

3.2 Cost recovery

With most scholarly publication and government-funded survey, the main costs are incurred in prepublication research. Even the costs of publication relate mostly to preparatory work before the first copy is printed (B 1). The effect of electronic delivery is to reduce the initial publication cost and almost eliminate the costs of supplying subsequent copies, as printing costs fall on the user. The costs that might eventually be recovered include digitizing and storing the information, a contribution to the cost of its acquisition, and the overhead cost of maintaining the system standards and metadata. As mentioned earlier (M 2.2), success is helped by dominating the chosen market. It is therefore important for charges to be kept as low as practicable with the aim of attracting the largest possible number of customers. Customers require comprehensive information, and a viable system will need rapid growth both in terms of number of customers and amount of information. A prolonged period of free access while the service is being established, followed by gradual introduction of charges, is the pattern followed, for example, by most electronic journals and newspapers. Their large capital investment has no short-term return.

Registration of users can enable a repository to identify customers, find out which areas are of most interest and keep customers informed of relevant developments. It also ensures that the user is aware of the terms and conditions of supplying the information. The casual or one-off user can be allowed to bypass much of the registration procedure. For organizations that are heavy users of the information, a monthly or annual invoice could be convenient, covering all staff from that organization. This could either be a flat rate at levels related to usage, or based on the total of list prices for all objects accessed. For the large user, fixed amounts are simpler, but the occasional user may prefer to pay per object, and ways of transferring small amounts of cash for such transactions are being developed. For sums of more than a few dollars, charge cards are a possible alternative. The latest news of **charging procedures** can be found on the Web (see, for example, Schutzer, 1996; Herzberg, 1998).

Incentives are the driving force of an information system. An obvious incentive is money - the metric of utility space. As a creature of market forces, money can help to balance supply and demand. As an appendage to tradable objects, it can encourage sharing, not hoarding, of information. For example, a repository might charge a fee to depositors of information in order to recover the cost of managing and storing the information. The user of the information might also have to pay, to recover costs of dissemination and to pass on a royalty to the depositor. Academic susceptibilities might prefer a subsidized repository with payment in kudos not cash. Either way, there are incentives for all concerned to behave in a socially desirable manner.

To ensure that authors can benefit from their creativity, the law recognizes **intellectual property rights** (**IPR**), such as **copyright**. This covers the author's rights to acknowledgment (paternity), to avoid alteration by others (integrity) and to royalties from sale of the work. The ease of copying electronic documents puts IPR at risk, see Lejeune (1999) or section 5.1: legal issues in Bailey (1996). This is one impediment to electronic communication. So-called trusted systems have been developed, but not yet widely adopted, which enable the information supplier to control information distribution as never before (The Economist, 1999). Another problem is the difficulty of calculating value. Devising a simple but effective pricing mechanism involves compromise. For example, consider what some economists call **network externalities**, that is, activities that support, benefit and extend the system as a whole, rather than individual users.

You may recall Mr Bell's problem. He invented and built a telephone, but had no-one to call. There is a snowball effect. The more people own phones, the more useful each one is, provided of course that they all follow suitable standards to make communication possible, and their phone numbers are widely known. There might be profit for Mr Bell in setting up a telephone company and selling services; but it is then to his advantage to encourage and subsidize the network of lines and exchanges, the availability of directories, and to reward the initial subscribers until the snowball effect takes over. These network externalities are a necessary development cost to him, not a profit. Above all, he must remain locked into the dominant standards. Someday his telephone might be linked to others throughout the world. I suppose he could have made an alternative decision to give away telephones and profit from the sale of directories, but the customer's perception of value and the difficulties of protecting market share must be taken into account.

Standards and metainformation, which describe what information is available, what it is useful for, how to get it and what it means, can be regarded as network externalities in the information system. They enhance the value of the main body of information. The more widely known and accepted they are, the more the overall value is enhanced. There is therefore a case for making metainformation readily and freely available to all, or even paying for its dissemination (advertising). It follows that standard setting bodies need external funding from members or governments.

Another quirk of the system reflects the difficulty of the first purchaser of a telephone. The high cost and unreliability of the untried device are matched by the tedium of being able to chat only to Mr Bell. Initial involvement with a radically new information system, as contributor or user, has similar drawbacks. Being a pioneer is a mug's game - much better to wait until the systems are grooved in and most information transformed. For the rational individual, the clever strategy appears to be to wait until the last minute before leaping onto a new development curve. And so, for a while, governments, not wishing to be bypassed by history, offer subsidies for new developments. Rational organizations grab them, for an organization changes more slowly than an individual, with more to gain by being ahead of the field. They invent ways to motivate staff and customers – and the attractions of the rational employee fade in comparison with the one with knowledge of IT.

5. References

Adlam, K.A.McL., Clayton, A.R., Kelk, B., 1988. A 'demonstrator' for the National Geosciences Data Index. International Journal of Geographical Information Systems, 2(2), 161-170.

Butler, D., 1999. The writing is on the web for science journals in print. Nature, 397 (6716), 195-200.

CCTA, 1989. The Information Systems Guides. John Wiley & Sons, Chichester.

Herzberg, F., Mausner, B., Snyderman, B.B., 1993. The Motivation to Work. Wiley, New York, 157pp.

Kent, W., 1978. Data and Reality. North-Holland Publishing Company, Amsterdam, 211pp.

Laxton, J.L., Becken, K., 1995. The design and implementation of a spatial database for the production of geological maps. Computers & Geosciences, 22(7), 723-733.

Newell, R.G., Theriault, D., Easterfield, M., 1992. Temporal GIS - modeling the evolution of spatial data in time. Computers & Geosciences, 18(4), 427-434.

Ovadia, D.C., Loudon, T.V., 1993. GIS in a geological survey's migration strategy. Proceedings of the 5th National AGI Conference, Birmingham, UK. pp. 3.12.1-3.12.4.

Peuquet, D.J., Bacastow, T., 1991. Organizational issues in the development of Geographical Information Systems: a case study of U.S. Army topographic information automation. International Journal of Geographical Information Systems, 5(3), 303-319.

The Economist, 1999. Digital rights and wrongs. The Economist, 353(8128) (July 17 1999), 99-100.

5.1 Internet references

Australian Geodynamics Cooperative Research Centre. 4D geodynamic model of Australia. http://www.agcrc.csiro.au/4dgm/

Bailey, C.W., Jr., 1996- . Scholarly electronic publishing bibliography. Houston: University of Houston Libraries, 1996-99. http://info.lib.uh.edu/sepb/sepb.html

BRGM. Le programme national de recherche scientifique pour l'imagerie géologique et géophysique de la France en 3D.
http://www.brgm.fr/geofrance3d/geofrance3d.html

Buehler, K., McKee, L. (editors), 1998. The OpenGIS guide: Introduction to Interoperable Geoprocessing. http://www.opengis.org/techno/guide.htm

Herzberg, A., 1998. Safeguarding digital library contents: charging for online content. D-Lib Magazine, January 1998.
http://www.dlib.org/dlib/january98/ibm/01herzberg.html

Institute of Physics, 1999. Sources, Journals. http://www.iop.org/jo.html

Lejeune, L., 1999. Who owns what? The Journal of Electronic Processing, March 1999, vol 4 (3). http://www.press.umich.edu/jep/04-03/glos0403.html

Lithoprobe: Canada's National Geoscience Project
http://www.geop.ubc.ca/Lithoprobe/public/aboutlp.html

Netscape, 1999. Building applications in the Net economy.
http://developer.netscape.com/docs/wpapers/index.html

Schutzer, D., 1996. A need for a common infrastructure: digital libraries and electronic commerce. D-Lib Magazine, April 1996. http://www.dlib.org/dlib/april96/04schutzer.html

ScienceDirect, 1999- . ScienceDirect: providing desktop access to the full text of more than 1000 scientific, medical and technical journals published by the world's leading scientific publishers. http://www.sciencedirect.com/

Disclaimer: The views expressed by the author are not necessarily those of the British Geological Survey or any other organization. I thank those providing examples, but should point out that the mention of proprietary products does not imply a recommendation or endorsement of the product.

<<<Back to Table of Contents

On to Part M(4): Epilog (or Prolog)>>>